

(Citation 3)

JP Patent Application Disclosure No.57-102699-June 25, 1982

Application No. 55-179945-December 18, 1980

Applicant: Matsushita Electric Industrial Co., Ltd, Tokyo (JP)

Title: VOICE RECOGNIZER

Detailed Description of the Invention:

.....

The switch (18) is connected to the side E at the registration time, and to the side R at the recognition time. The switch (19) controls as to whether the signal line (20) is transmitted to the switch (21), and is connected at the registration time. If constituted in this manner, when the voice word is generated twice at the registration time, the voice of the first time is stored in the registration pattern area (I) (17), and the voice of the second time is stored in the input pattern area (3), and the pattern matching unit (6) can calculate a degree of similarity with the voice of the first time and the voice of the second time. Although the judging unit (7) performs the same process as the judging unit of Fig. 4, at this time, the degree of similarity with the input pattern is calculated only with the registration pattern area (I), and hence, the Smax becomes the

degree of similarity with the registration pattern area (I) and the input pattern.

If the formula (1) is satisfied, the judging unit (7) allows the switch (21) to turn ON through the output line (20), and stores the content of the registration pattern area (I)(17) in the registration pattern area (II)(5) together with the category code of the input terminal (8). The registration pattern area (II)(5) is an area which stores the voice word to be recognized similarly to the registration pattern area of Fig. 4. If the formula (1) is not satisfied, the switch (21) is caused to turn ON by the output line (20). At the same time, whether the voice word is registered in the speaker by the output (9) is outputted. When not registered, the speaker further pronounces the same voice word. At this time, though it is the pronunciation of the third time, the degree of similarity with the pattern corresponding to the pronunciation of the second time and the pattern corresponding to the pronunciation of the third time is similarly calculated, and this is judged by the judging unit (7), and is subjected to the above described process. In this manner, until the degree of similarity becomes larger than a threshold value K, the same voice is pronounced as many times as necessary. Only when the degree of similarity becomes larger than the threshold value K, the registration pattern area (II)(5) is stored with the register

pattern, and after all, the registration pattern area is stored with a regular pattern, and an irregular pattern, for example, the pattern mixed with ambient noise is not registered.

.....

⑯ 日本国特許庁 (JP)

⑰ 特許出願公開

⑰ 公開特許公報 (A)

昭57-102699

⑮ Int. Cl.<sup>3</sup>  
G 10 L 1/00

識別記号

序内整理番号  
7350-5D

⑯ 公開 昭和57年(1982)6月25日

発明の数 1  
審査請求 未請求

(全 6 頁)

④ 音声認識装置

⑦ 発明者 藤内秀和

門真市大字門真1006番地松下電

器産業株式会社内

⑧ 特願 昭55-179945

⑨ 出願 昭55(1980)12月18日

⑩ 出願人 松下電器産業株式会社

⑪ 発明者 大賀英文

門真市大字門真1006番地

松下電器産業株式会社内

⑫ 代理人 弁理士 大島一公

明細書

1. 発明の名称

音声認識装置

2. 特許請求の範囲

① 音声登録バタンエリアと音声入力バタンエリアとを設け、音声登録バタンエリアにはあらかじめ認識すべき音声単語に対応する音声バタンを格納しておいて、前記登録バタンエリア内の音声バタンと、前記音声入力バタンエリア内の音声バタンとのバタンマッチングを行なうことによつて入力バタンエリア内の音声バタンに対応する音声単語を認識する音声認識装置において、登録すべき音声は2回以上発声することにし、前記入力バタンエリアに新たなバタンが入力される毎に、以前の入力バタンエリア内のバタンを退避する登録バタンエリアを設け、入力バタンエリア内の音声バタンとのバタンマッチングを行なつて類似度を算出し、その類似度が予め定められたシキイ値より大きい時にのみ前記登録バタンエリア(1)内の音声バタンあ

るいは前記入力バタンエリア内の音声バタンを登録バタンエリア(1)へ転送することを特徴とする音声認識装置。

② 入力バタンエリア内の音声バタンと以前の入力バタンエリア内のバタンを退避する登録バタンエリア内の音声バタンとのバタンマッチングを行ない、この時の類似度をS<sub>1</sub>とし、さらに前記入力バタンエリア内の音声バタンと、前記登録バタンエリア内の音声バタンとのバタンマッチングを行ない、最大の類似度を抽出しこれをS<sub>2</sub>として、前記S<sub>1</sub>がシキイ値より大でかつS<sub>1</sub>がS<sub>2</sub>より大である時のみ、前記バタンを退避する登録バタンエリア内のバタンを登録バタンとして登録バタンエリアへ転送することにした特許請求の範囲第1項記載の音声認識装置。

3. 発明の詳細な説明

本発明は、あらかじめ登録された登録バタン群と入力バタン間とのバタンマッチングによつて、入力バタンが登録バタン群のどのカテゴリに最も似ているかを判断し、入力バタンを識別するバタ

ンマッチング識別方法を音声に利用した音声認識装置に関するもので、その目的とするところは登録パタンを登録する場合に正当でないものを排除するよう制御を加えようとするものである。

パタンマッチング方式における音声認識方法について第4図に従つて説明する。発声は孤立音声単語であり、音声はマイク①によって電気信号に変換され、特徴抽出部②によつて特徴が抽出されて特徴パタンに変換される。又特徴抽出部②では音声区間の検出がなされ、発声された孤立音声単語に対応した特徴パタンは入力パタンエリア④に格納される。④は認識すべき音声単語の登録パタンとそのカテゴリコードを格納している登録パタンエリアであり、スイッチ⑥はE(ENTRY) モードと接続している場合には入力パタンエリア④に格納されている特徴パタンは入力端子⑨より入力されているカテゴリコードとともに登録パタンエリア④へ転送される。あらかじめ、スイッチ⑥をDモードにして、認識すべき音声単語を発声して、その特徴パタンと、カテゴリコードを登録エリアへ格納す

(3)

なされてリジエクトされ、リジエクトに対応したコードが outputされる。以上の処理は判定部⑤で行なわれる。

このような音声認識装置においては、登録に関して正当でない登録パタンがそのまま登録されてしまうという問題があつた。登録すべき音声単語を発声した時に、たまたま雜音が加わつて来た場合、あるいはノードの調子がおかしい時に発声した場合等に得られた登録パタンはかららざるも正当なものではない。正当でない登録パタンのままで認識されれば、結果として誤認識が多発することとなる。

本発明は、登録されるパタンが正当なものか否かを判断し、正当なものである場合には登録し、正当なものでない場合には登録しないようにすることにより、從来より問題であつた正当でない登録パタンによる誤認識を減少させたものである。

以下、実施例として示した図面に従つて、その構成を説明する。

①のマイク、②の特徴抽出部、④の入力パタン

る。④は入力パタンエリア④の入力パタンと登録パタンエリア④内の複数の登録パタンとのパタンマッチングを行ない、登録パタン毎に入力パタンとの類似度を出力線図を介して出力するパタンマッチング部である。この約には、スイッチ⑥はE(Recognition) モードと接続される。④はパタンマッチング部④からの出力である各登録パタン毎の類似度を出力線図で受けるとともに、その時のカテゴリコードを出力端子より受けて、入力パタンがどの登録パタンと最も似ているかを判別する判定部である。登録パタン毎の類似度の中から最も大きい(最も類似性のある)値  $s_{max}$  を選出する。この  $s_{max}$  があらかじめ定められたシキイ值  $K$  より大なる時、

$$\text{つまり}, \quad s_{max} > K \quad \cdots \cdots \text{④式}$$

の時に入力パタンは  $s_{max}$  を出力した登録パタンであると識別され、その登録パタンのカテゴリコードが出力される。 $s_{max}$  が④式を満足しない時は登録パタンエリア内に格納されている登録パタンとは異なるたつ入力パタンが入力されたとみ

(4)

エリア、④の登録パタンエリアおよび④のパタンマッチング部について第4図と同様な構成で、同じ動作である。

登録時には、同じ音声単語を2回以上発声することとし、カウンタ部は何回同一のカテゴリコードを有する音声単語が発声されたかをカウントするものである。特徴抽出部において、音声区間の検出はなされるが、音声単語の終端が検出されるると出力端子を介してクロックを出し、カウンタ部はそのクロックによつてカウントアップする。すなわち、音声単語を発声するたびにカウントアップされる。一方カウンタ部はリセット時換出部⑤からの信号によつてリセットされる。リセット時換出部⑤は入力端子⑨より入力されるカテゴリコード(登録されるべき音声単語に対応している)を受けて、そのカテゴリコードが変化したか否かを検出し、カテゴリコードが変化した場合にはカウンタ部をリセットする。同じ音声単語であるなら当然そのカテゴリコードも同一であるため、このような構成によつて登録する音声単語が何回発

(5)

(6)

両されたかを知ることが出来る。

カウンタ時<sup>(1)</sup>の出力時はスイッチ時<sup>(2)</sup>を制御し、登録すべき音声単語の発声が第1回目であるなら、スイッチ時<sup>(2)</sup>をオフし、2回目以降ならスイッチ時はオンさせるように制御する。入力端子<sup>(3)</sup>のカティゴリコードを変化させない限りスイッチ時はオン状態となるため、認識時にカティゴリコードを変化させないようにすればスイッチ時はオン状態になる。認識時にはスイッチ時はオン状態にするように出力線<sup>(4)</sup>を制御しても良い。

登録バタンエリア<sup>(1)</sup>時の入力は、入力バタンエリア<sup>(3)</sup>の出力と接続され、特徴抽出部<sup>(4)</sup>からの音声バタンが入力バタンエリア<sup>(1)</sup>へ格納される時に、入力バタンエリア<sup>(1)</sup>に格納されていた以前のバタンは登録バタンエリア<sup>(1)</sup>時<sup>(2)</sup>に転送される。従つて、登録バタンエリア<sup>(1)</sup>時の内容は、1つ前に発音された音声単語の特徴バタンを格納していることとなる。

スイッチ時は登録時にはR側に、認識時にはR側に接続される。スイッチ時は信号線<sup>(5)</sup>をスイッ

(7)

(b) によって話者に登録されたか否かを出力する。登録されない場合には、さらに話者は同一の音声単語を発音する。この時には3回目の発声となるが2回目の発声に対応したバタンと3回目の発声に対応したバタンとの類似度が同様算出され、判定部<sup>(6)</sup>で判定されて、前述したような処理がなされる。このようにすることによって、類似度がシキイ値<sup>(7)</sup>より大きくなるまで、何回でも同一の音声を発声することとなる。シキイ値<sup>(7)</sup>より大になれた時にはじめて登録バタンエリア<sup>(1)</sup>には登録バタンが格納されることとなり、結局登録バタンエリアには正当なバタンが格納され、正当でないバタン、例えば周囲雜音が混在したようなバタンは登録されないととなる。

認識時には、スイッチ時をR側に接続しておき、かつスイッチ時<sup>(2)</sup>をオフしておいて、スイッチ時<sup>(2)</sup>が常にオフするようにしておけば良い。登録バタンエリア<sup>(1)</sup>と入力バタンエリア<sup>(3)</sup>とのバタンマッチングをバタンマッチング部<sup>(4)</sup>で行なつて、判定部<sup>(6)</sup>では前述したと同様な処理を行なえば、入力

テ御へ伝達するか否かを制御するもので、登録時には接続している。このような構成にしておけば登録時に音声単語を2回発生すれば、1回目の音声は登録バタンエリア<sup>(1)</sup>時<sup>(2)</sup>へ格納され、2回目の音声は入力バタンエリア<sup>(3)</sup>時<sup>(2)</sup>へ格納されることとなり、バタンマッチング部<sup>(4)</sup>は第1回目の音声と、第2回目の音声との類似度を算出することができる。判定部<sup>(6)</sup>は、第4図の判定部と同様な処理を行なうが、この時には登録バタンエリア<sup>(1)</sup>のみとしか入力バタンとの類似度は算出されないため、前記したSmax<sup>(8)</sup>は、登録バタンエリア<sup>(1)</sup>と入力バタンとの類似度となる。

式を満足しておれば、判定部<sup>(6)</sup>は出力線<sup>(4)</sup>を介してスイッチ時<sup>(2)</sup>をオンさせ、入力端子<sup>(3)</sup>のカティゴリコードとともに、登録バタンエリア<sup>(1)</sup>時の内容を登録バタンエリア<sup>(1)</sup>時<sup>(2)</sup>へ格納する。登録バタンエリア<sup>(1)</sup>時は、第4図の登録バタンエリアと同様認識されるべき音声単語を格納しているエリアである。もし式を満足していなければ出力線<sup>(4)</sup>によつてスイッチ時<sup>(2)</sup>をオンさせる。又同時に出力

(8)

バタンの識別が可能となる。始め登録バタンエリア<sup>(1)</sup>からの出力線である。登録バタンエリア<sup>(1)</sup>には正当なバタンしか格納されてないため、従来に比べて認識率は良くなり、誤認識も減少することについた。

第2図は、判定部をさらに改良したものである。動作としては判定部<sup>(6)</sup>と判定部<sup>(6)</sup>は同様であるが、認識時と登録時でシキイ値<sup>(7)</sup>を変化させる。スイッチ時は認識時にはR側に、登録時にはR側に接続され、認識時にはシキイ値<sup>(7)</sup>はRIC<sup>(9)</sup>登録時にはKC<sup>(10)</sup>になる。

$$K_E > K_R$$

としておくことにより、登録時にはさらにきびしく登録バタンが正当か、否か判別されることになる。登録バタンエリアには、より正当性のあるバタンしか登録されないことになる。このようにすればさらに認識率の向上、誤認識の減少が可能になる。

又、第8図に示すような構成も考えられる。入力バタンマッチング部<sup>(4)</sup>は、入力バタンエリア<sup>(3)</sup>

の内容と登録バタンエリア(I)回の内容とは、認識時登録時に関係なくバタンマッチングを行なうようにし、登録バタンエリア(I)回とは、登録時のみ入力バタンとマッチングを計算するようにする。スイッチ回は從つて登録時のミオンシ、認識時はオフしている。さらに判定部回へは登録時のカテゴリコードを示している信号線回を登録時のみ加わる様な構成にしておく。スイッチ回は、登録時のみオンするようにしておく。カテゴリコード判定部回は出力線回の内容と信号線回の内容を比較し、等しい時のスイッチ回をオンするように制御線回を介してスイッチ回を制御する。

本発明はこのような構成をとつたので、登録時にはこれから登録しようとするバタンの龜にすでに登録されているバタンともふくめて、バタンマッチングを行なうことができ、従つて判定部回で選出される最も類似度の高い値  $S_{max}$  は、すでに登録されているバタンにおける類似度もふくめた値になり、この  $S_{max}$  を出力したカテゴリコードが 出力線回に出力される。このコードがこれから

00

又、発声者は出力回の内容から登録されたか否かを判断することができる。指定したカテゴリコードが 出力線回に出力された場合には、登録されたことになり、逆に 出力時のコードと指定したコードとが異なる場合には登録されないと示す。出力回の内容が信号線回の内容と異なる場合には登録しようとする音声単語がすでに登録されている音声単語のどれと近いかを知ることができる。従つて、登録しようとする音声単語を変更する等の処理をすれば良い。例えはイチ(I)とシチ(I)ならシチをナナと發音するように変更する。

さらに判定部の処理を次に示すようにすれば、上記のような似かよつたバタン同士の登録を防ぐことができる。バタンマッチング部から出力される類似度の中から 2 番目に高い類似度を  $S_{max2}$  とすると

$$S_{max} - S_{max2} > L \quad \cdots \text{回式}$$

という条件を加えれば良い。このようにすれば、すでに登録されているバタンとすくなくとも類似度が L 以上のものしか登録されないことになる。

登録しようとするカテゴリコードであつた場合にはカテゴリコード判定部回によつてスイッチ回をオンし、登録バタンエリア(I)回の内容を登録バタンエリア(I)回へ転送するように出力線回に出力する。 $S_{max}$  を出力したカテゴリコードがこれから登録しようとするカテゴリコードでなかつた場合にはカテゴリコード判定部回によつてスイッチ回をオフするように出力線回に出力する。

又、最も高い類似度  $S_{max}$  が K より小の場合には、前述したと同様リジエクトに対応したコードが出力線回に出力され、当然これから登録しようとするカテゴリコードとは異なるためスイッチ回のオンはなされない。これから登録しようとするバタンが、すでに登録されているバタンの中に良く似ているバタンである場合(例えは、オカヤマとワカヤマ、イチ(I)とシチ(I)、ナハとナラ)にはすでに登録されているバタンの類似度の方がなんらかの原因(発音の不安定さ、あるいは周囲のノイズ)で、高いことがあります。第 8 図のようにしておけばこのような場合には登録されないことになる。

01

従つて誤認識はさらに改良される。

又、第 1 図、第 8 図においては登録バタンエリア(I)回を設けたがこれを設けず、登録バタンエリアの内容の一部を消去する機能を設けても良い。登録時には登録エリアへ信号を転送するようにしておき、登録時の判定の結果、登録してはいけない場合と判明した時に限つて、登録カテゴリコードに対応した登録バタンの内容を消去するようすれば第 1 図あるいは第 8 図と同様な処理が可能となる。又、第 1 図において、登録時に正当なバタンであると判断された時には、登録バタンエリア(I)回内のバタンを登録バタンエリア(I)回へ転送せず、入力バタンエリア回の内容を登録バタンエリア(I)回へ転送するようにしても良い。

本発明は上記のように稚育、あるいは発声の不安定さ等による原因で生じた正当でないバタンは登録されないために、誤認識を減少することができ、登録時に制限を加えることにより、似かよつたバタンの登録を防ぐことができ、安定した認識ができる効果がある。又、判定部のシキイ値を登

02

03

録時には認識時よりきびしくする方向にすれば、  
より正当なパタンしか登録されないため、誤認識  
がさらに減少する等の効果が得られる。

#### 4. 図面の簡単な説明

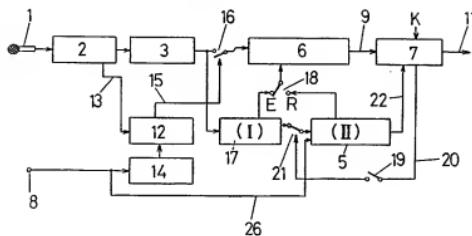
第1図は本発明装置の構成図、第2図は判定部  
の検査施設を示す構成図、第3図は本発明装置の  
検査施設を示す構成図、第4図はパタンマッチン  
グ方法による音声認識装置の構成図。

- (a) - 入力パタンエリア、(b) - 登録パタンエリア  
(I)、(II) - パタンマッチング部、(d) - 判定部、(e)  
- 登録パタンエリア (I)

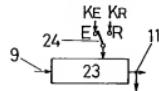
代理人 弁理士 大島一公

等

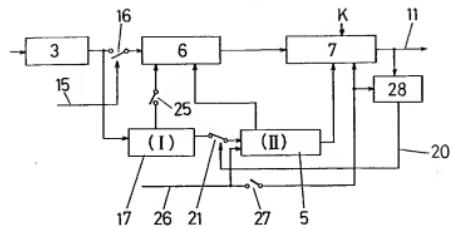
第1図



第2図



第3図



第4図

